

D9.4 Database with the metadata of surveys to EMMs across Europe

Ami Katherine Saji, Laura Morales

▶ **To cite this version:**

Ami Katherine Saji, Laura Morales. D9.4 Database with the metadata of surveys to EMMs across Europe. Sciences Po, CEE. 2020, 29 p. hal-04001050

HAL Id: hal-04001050
https://sciencespo.hal.science/hal-04001050

Submitted on 22 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution| 4.0 International License

Ref. Ares(2020)5455326 - 13/10/2020

Research and Innovation Action

Social Sciences & Humanities Open Cloud

Project Number: 823782 Start Date of Project: 01/01/2019 Duration: 40 months

Deliverable 9.4 Database with the metadata of surveys to EMMs across Europe

Dissemination Level	PU
Due Date of Deliverable	31/10/20 (M22)
Actual Submission Date	13/10/20
Work Package	WP9 Data Communities
Task	T9.2 Ethnic and Migration Studies
Type	Report
Approval Status	Approved by EC - 03 Nov 2020
Version	V1.0
Number of Pages	p.1 – p.29

Abstract:
This report describes how the FAIR principles have informed the design, development, and delivery of the Ethnic and Migrant Minorities (EMM) Survey Registry – an online and public database and tool that displays compiled survey-level metadata of existing quantitative surveys conducted with EMM populations in Europe and which is being produced in the Social Sciences and Humanities Open Cloud project (SSHOC) in close collaboration with the COST Action 16111 – ETHMIGSURVEYDATA and the French Agence Nationale de la Recherche (ANR)-funded project, FAIRETHMIGQUANT.

The information in this document reflects only the author's views and the European Community is not liable for any use that may be made of the information contained therein. The information in this document is provided "as is" without guarantee or warranty of any kind, express or implied, including but not limited to the fitness of the information for a particular purpose. The user thereof uses the information at his/ her sole risk and liability. This deliverable is licensed under a Creative Commons Attribution 4.0 International License.

SSHOC, "Social Sciences and Humanities Open Cloud", has received funding from the EU Horizon 2020 Research and Innovation Programme (2014-2020); H2020-INFRAEOSC-04-2018, under the agreement No. 823782

History

Version	Date	Reason	Revised by
0.0	10/09/2020	Initial outline of the report	Laura Morales, Ami Saji
0.1	15/09/2020	First draft based on report outline	Ami Saji
0.2	18/09/2020	Review of V0.1 by T9.2 lead. Revision of V0.1 based on review feedback	Laura Morales, Ami Saji
0.3	25/09/2020	Peer review of V0.2. Revision of V0.2 based on review feedback	Cees Van Der Eijk, Ami Saji
0.4	04/10/2020	Final revisions after editorial review by CO	Laura Morales, Ami Saji
1.0	13/10/2020	Final version for submission	Laura Morales, Ami Saji

Author List

Organisation	Name	Contact Information
Sciences Po	Ami Saji	amikatherine.saji@sciencespo.fr
Sciences Po	Laura Morales	laura.morales@sciencespo.fr

Time Schedule before Delivery

Next Action	Deadline	Care of
Initial outline of the report	10/09/2020	Sciences Po
Drafting and inputting of text based on report outline	15/09/2020	Sciences Po
Review of V0.1 by T9.2 lead	17/09/2020	Sciences Po
Revision of V0.1 based on review feedback	18/09/2020	Sciences Po
Peer review of V0.2	24/09/2020	UNOTT
Revision of V0.2 based on review feedback	25/09/2020	Sciences Po
Review of V0.3 by SSHOC project coordinators	30/09/2020	CESSDA
Revision of V0.3 based on review feedback	04/10/2020	Sciences Po
Final review and submission	13/10/2020	Sciences Po, CESSDA

Executive Summary

This document is a deliverable of the Social Sciences and Humanities Open Cloud project (SSHOC). It is a report about the Ethnic and Migrant Minorities (EMM) Survey Registry[1]: an online, publicly available database and tool that displays compiled survey-level metadata of existing quantitative surveys undertaken with ethnic and migrant minority (EMM) populations in Europe, and which is jointly being produced by SSHOC's Task 9.2 *Ethnic and Migration Studies* (part of Work Package 9 *Data Communities*) in collaboration with COST Action 16111 – ETHMIGSURVEYDATA[2] (a network of more than 200 European researchers active in the ethnic and migration studies field) and FAIRETHMIGQUANT[3] (a French Agence Nationale de la Recherche (ANR)-funded Open Science project). This report is intended to inform prospective users of the EMM Survey Registry (i.e. (non)academic researchers, policymakers, and other practitioners interested in quantitative surveys on EMMs' integration) about what the registry is and how it could be leveraged for their own research and/or policy-related activities. It is also aimed at providing individuals and organisations involved in data curation with a concrete example of how a FAIR-friendly tool for the social sciences can be created from a bottom up approach, where the data community is the main driver.

The report begins by illustrating the role of the FAIR (Findable, Accessible, Interoperable, Reusable) principles in shaping how the EMM Survey Registry has been designed, developed, and delivered. The EMM Survey Registry is currently available in 'beta' version, and information for over 400 surveys from 11 different countries are publicly being displayed. More specifically the report follows the trajectory of the EMM Survey Registry, describing how stakeholders specialising in data production, usage, storage, and management were routinely leveraged from conception to delivery (of its current 'beta' version), in order to align this registry with the FAIR principles as much as possible. The report concludes by describing the concerted effort the Task 9.2 team is making to facilitate short and long-term sustainability of the EMM Survey Registry by strategically connecting and liaising with potential users interested in quantitative surveys undertaken with EMMs.

[1] The Ethnic and Migrant Minorities (EMM) Survey Registry: https://ethmigsurveydatahub.eu/emmregistry/; [1 Oct 2020]
[2] The COST page about COST Action 16111 ETHMIGSURVEYDATA: https://www.cost.eu/actions/CA16111/; [1 Oct 2020]
[3] A description about FAIRETHMIGQUANT provided on the Sciences Po website: https://www.sciencespo.fr/centre-etudes-europeennes/en/node/30931; [1 Oct 2020]

Abbreviations and Acronyms

ANR	Agence nationale de la recherche
API	Application Programming Interface
CESSDA	Consortium of Social Sciences Data Archives
CDSP	Centre de Données Socio-Politiques, Sciences Po
DDI	Data Documentation Initiative
EMM	Ethnic and migrant minority
EOSC	European Open Science Cloud
EQB	European Question Bank
ETHMIGSURVEYDATA	The International Ethnic and Immigrant Minorities' Survey Data Network
FAIR	Findable, Accessible, Interoperable, Reusable
FAIRETHMIGQUANT	Making Ethnic and Migrant Minority Survey Data FAIR
GESIS	Leibniz Institute for the Social Sciences
OAI-PMH	Open Archives Initiative Protocol for Metadata Harvesting
QDB	Question Data Bank
Sciences Po	Fondation Nationale des Sciences Politiques
SSHOC	Social Sciences and Humanities Open Cloud
T	Task
UNOTT	University of Nottingham
WP	Work Package
XML	Extensible Markup Language

Table of Contents

1. Introduction ***6***

2. Designing and Developing the EMM Survey Registry ***8***

2.1 Addressing a Community Need for a FAIR-Compliant Tool **8**

2.2 Enhancing User Experience of the EMM Survey Registry **9**

2.2.1 Creating and Using a Rich and Meaningful Metadata Schema 9

2.2.2 Centering the Front-End on the User 12

2.2.3 Managing the EMM Survey Registry through the Back-End 16

2.3 Adopting DDI Codebook **19**

2.4 Setting up the EMM Survey Registry to be a Sustainable Tool **21**

3. Connecting and Liaising with Prospective Users of the EMM Survey Registry ***23***

4. Conclusion ***27***

5. References ***28***

6. List of Images ***29***

7. List of Tables ***29***

1. Introduction

As part of Work Package 9 *Data Communities* of the SSHOC project, Task 9.2 *Ethnic and Migration Studies* brings together the data community focused on ethnic and migration studies. Its overarching objective is to make quantitative survey data on ethnic and migrant minorities (EMMs) more accessible and (re)usable to a wide range of users (e.g. (non)academic researchers, policy makers) in Europe and beyond. As a means to achieve this objective, the Task 9.2 team has planned for two tangible outputs: first, launching a free online database and tool that will be FAIR (Findable, Accessible, Interoperable, Reusable)-compliant[4] and will display compiled survey-level metadata for over 800 quantitative EMM surveys from over 30 different European countries[5], i.e. the EMM Survey Registry; and second, testing the feasibility of setting up, as part of the CESSDA-led European Question Bank (EQB), a component dedicated to EMM surveys identified via the development of the EMM Survey Registry.

The SSHOC project is thus providing the funding and technical support needed by the Task 9.2 team to successfully deliver these two outputs. Moreover, as Task 9.2 pertains to the data community for ethnic and migration studies, it involves the participation and support of two different initiatives also invested in making EMM survey data FAIR: COST Action 16111 – ETHMIGSURVEYDATA and FAIRETHMIGQUANT. The former is an international network funded by the COST Association with more than 200 EMM-focused researchers primarily based across 35 countries in Europe and neighbouring countries and it provides the intellectual impetus for the two outputs[6]. The latter is an Open Science project funded by the French Agence Nationale de la Recherche (ANR) and is responsible for ensuring the inclusion of the French EMM surveys in the two outputs, as well as promoting the FAIR principles with the ethnic and migration studies data community in France.

The first output, the production of the EMM Survey Registry, had foreseen a target completion date of October 2020. Having successfully met this target, the Task 9.2 team has produced this report to document the current state (i.e. the publicly available 'beta' version with records for over 400 surveys from 11 different European

[4] The guidance provided to H2020 beneficiaries by the European Commission's Directorate-General for Research & Innovation on how to make research data FAIR: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf; [1 October 2020]

[5] The 30 plus countries that will be contributing to the EMM Survey Registry are those that have formally joined ETHMIGSURVEYDATA. To date, there are 35 countries that are participating in ETHMIGSURVEYDATA and the full list can be accessed here: https://www.cost.eu/actions/CA16111; [1 October 2020]. For our purposes, we refer to all the countries participating in ETHMIGSURVEYDATA as 'European' as they are participating in a European-led or focused research network; however, it is important to note that the 35 countries include Turkey and Israel.

[6] COST Actions are research networks funded by COST (European Cooperation in Science and Technology). They are intended to foster research collaboration and innovation through networking activities (e.g. meetings, conferences, training schools, short-term research stays, publications, dissemination activities). Countries (primarily those from Europe) formally join COST Actions to support their respective objectives and goals. COST Actions do not fund the actual costs of undertaking research (human resources or technical/technological development).

countries) of the EMM Survey Registry. It specifically describes how the EMM Survey Registry was purposefully designed, developed, and delivered to align with the FAIR principles and how the ethnic and migration studies data community[7] (i.e. the Task 9.2 team alongside ETHMIGSURVEYDATA and FAIRETHMIGQUANT) has already been able to promote short and long-term sustainability of the registry through strategic dissemination and outreach.

This report is intended to provide a detailed overview of the present status of the EMM Survey Registry, as well as the correlated work undertaken by the data community. As such, it is structured into the following three sections:

- **Designing and Developing the EMM Survey Registry:** A summary of the decisions made by the data community to design and develop the EMM Survey Registry's metadata schema and its front and back-ends to be aligned as much as possible with the FAIR principles. Also, a discussion of the actions taken by the data community to ensure that the EMM Survey Registry is embedded into a sustainable infrastructure that will, in turn, help ensure the registry's relevance and usability in the short and long-term.
- **Connecting and Liaising with Prospective Users of the EMM Survey Registry:** A description of the diverse outreach and dissemination activities taken up by the data community to gain support from prospective users of the EMM Survey Registry in fostering the short and long-term sustainability of the tool.
- **Conclusion:** A reflection on the important role the EMM Survey Registry can play in helping to make quantitative survey data on EMMs' integration FAIR.

In reading this report, it is important to note that the current state of the EMM Survey Registry is not the final version. As mentioned above, the 'beta' version, which is already publicly available at the time of writing, only displays information for 435 of the 800 plus surveys that have been detected in the 30 different European countries. Moreover, while the 'beta' version has all of the planned functionalities for both the front-end (i.e. the interface that allows users to explore the metadata that has been compiled for the different surveys) and back-end (i.e. the interface that allows for all the EMM Survey Registry's (meta)data to be stored, organised, and controlled), new opportunities to improve the registry's user-friendliness have been identified, in the short, medium and long-term. Thus, the current 'beta' version status means that there will be future improved versions of the EMM Survey Registry, and the data community is actively working to map out what these future versions will look like. To date, the firm confirmation that can be made is that a future final version will include the full 'census' of surveys identified by the data community for around 30 European countries.

[7] Throughout this report the term, 'the data community' is used to refer to the one for ethnic and migration studies, which consists of the Task 9.2 team plus the two correlated projects, ETHMIGSURVEYDATA and FAIRETHMIGQUANT.

2. Designing and Developing the EMM Survey Registry

2.1 Addressing a Community Need for a FAIR-Compliant Tool

European societies are not socially homogeneous, as many have been formed by a range of groups with linguistic, religious, national or racial elements of diversity; some of these groups have become dominant and synonymous with the nation state, while others have become minorities (often recognised as national or ethnic minorities). Europe has also, over the past 30 years, become an important destination for people migrating as a means to improve their life conditions or to escape from famine, persecution, or war (Morales et al., 2020b). With this ever-growing diversity of people calling Europe their home, EU institutions and member states have inevitably ramped up efforts to obtain informative and reliable indicators of integration for migrants and ethnic minorities (Huddleston et al., 2013).

ETHMIGSURVEYDATA, as a European research network with over 200 members from the ethnic and migration studies field, recognised the important and imperative role quantitative surveys play in studying and learning about the integration experiences of EMMs in Europe. At the same time, they were aware that such data was not always easy to find, access, and reuse, as they were not systematically being stored, documented, and shared.

Thus, herein lies the rationale behind building the EMM Survey Registry, which is now being developed by the data community for ethnic and migration studies (i.e. the Task 9.2 team plus the projects, ETHMIGSURVEYDATA and FAIRETHMIGQUANT). By designing and developing a tool that would be shaped by the FAIR principles, it would allow a wide range of users to untap the research potential of quantitative surveys undertaken with EMM (sub)populations. In other words, the EMM Survey Registry would apply the FAIR principles as follows:

- **Findable:** Quantitative surveys on EMMs' integration would be easier to locate, as the EMM Survey Registry would serve as a 'census' of existing surveys in Europe. It would be designed and developed to allow users to easily navigate and explore information about each survey, by offering user-friendly and user-centric interfaces on the front and back-ends of the registry and by systematically documenting information about each survey via a metadata schema, i.e. a set of metadata variables that each survey will have information coded for.
- **Accessible:** The EMM Survey Registry would be made publicly available online, ensuring that the information it houses is accessible to a wide range of users. Moreover, the EMM Survey Registry's metadata schema would include information about how to access the survey's data set(s), technical documentation, questionnaire(s), and other relevant publications.
- **Interoperable:** All the metadata housed on the EMM Survey Registry would be set up so that they could easily be shared with humans and machines (namely other social sciences data archives and repositories). At a minimum this has meant setting up the front-end interface so that the metadata is

presented in an easy to read and navigate format and is offered as XML files compliant with DDI Codebook. In the best-case scenario, this would mean also offering the DDI Codebook compliant XML files through an API that conforms with Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). *[NOTE: See section 2.3 below for more information about how the EMM Survey Registry is adopting DDI Codebook and where the EMM Survey Registry is in terms of offering the DDI Codebook-compliant XML files.]*

- **Reusable:** The metadata offered by the EMM Survey Registry would be detailed, informative, organised, and structured; as a result, users would be able to utilise the metadata itself for analysing and learning about quantitative surveys on EMMs' integration. The EMM Survey Registry is also promoting or facilitating reuse of the original survey as it offers information about how to access all relevant sources (i.e. datasets, technical documentation, questionnaires, publications) whenever possible[8].

As the EMM Survey Registry is driven by the data community with intimate collaboration with and support from data distributors/managers, it has made immense progress since it was initially conceptualised by the researchers participating in the ETHMIGSURVEYDATA network. To date, the current 'beta' version of the EMM Survey Registry was launched in spring 2020 and it offers users a fully functional front-end (see section 2.2 for more details) and access to information about over 400 different surveys from 11 different European countries (including DDI Codebook-compliant XML files). The data community has also conducted rigorous quality checks of metadata compiled for roughly 20 other European countries (an estimated 400-500 more surveys to be described and documented in the EMM Survey Registry), as well as explored different options to enhance the current offerings of the registry. These continued efforts by the data community are thus demonstrative of their desire of delivering, as the end product, a FAIR-compliant EMM Survey Registry.

2.2 Enhancing User Experience of the EMM Survey Registry

2.2.1 Creating and Using a Rich and Meaningful Metadata Schema

In order to entice users to the EMM Survey Registry, the data community determined that rich and meaningful survey-level metadata would need to be produced and offered in a user-friendly and DDI Codebook-compliant manner (see section 2.3 below about the adoption of DDI Codebook). As such, the data community employed an iterative process for developing the EMM Survey Registry's metadata schema, where each version of the metadata schema was rigorously discussed and tested by different academic and non-academic users of quantitative survey data on EMMs' integration. The final version, which exists as an Excel-based template[9] and

[8] Survey producers cannot deposit their data, technical documentation, questionnaire, and/or other relevant publications directly to the EMM Survey Registry. Therefore, the EMM Survey Registry can only facilitate reuse if the survey data, technical documentation, questionnaire, and other relevant publications have been made publicly available elsewhere (typically through a data archive or a research project website).

[9] The Excel-based template was produced by the data community prior to the start of the SSHOC project and as part of the research networking efforts funded by COST Action 16111- ETHMIGSURVEYDATA. It is now publicly available for

is organised into sections, captures pertinent information such as: the key features of the survey; the (EMM) target population of the survey; the sampling methods used; the sample sizes for the survey as a whole and for any partitioned subgroups; the method of data collection used; the availability of the survey data set(s), questionnaire(s), and technical documentation; and details and specifications about data ownership and distribution. The full list of information captured by the final version of the metadata schema is presented in Table 1 below.

Table 1: Metadata Schema by Section

Section	Information Covered in the Section
1. General Identification Information about the Survey	Full survey name (in English and in the native language), survey acronym, territorial scope of the survey (national vs. subnational), representativeness of the target population of the survey, type of survey (e.g. cross-section, longitudinal), survey start/end dates, main topics of the survey, main purpose of the survey, coverage of the survey's target population in terms of age, coverage of the survey's target population in terms of the sex of the respondents
2. Information about the Inclusion of the Survey in a Larger Study	Full name (in English and in the native language) of the larger study that the survey belongs to; acronym of the larger study; names of other countries/regions/cities (in English) that are also part of the larger study; for larger studies that are repeated cross-section/longitudinal, date of the first survey, frequency of the waves/panels, and wave number of the survey (since each wave is coded as its own record); for larger studies that are part of an international survey programme, date when the survey became a part of it and frequency of the waves since the survey joined it; for larger studies that have pooled samples, number of surveys pooled, other surveys with which the survey has been pooled, and indication as to whether or not emigrants from more than one country have been pooled; details about any qualitative studies that have been linked to the survey
3. EMM Target Population	Description of the survey's EMM target population, classification of the survey's EMM target population into type, ways in which the EMM target population has been operationalized (by the survey producer), migrant/minority-related questions that have been included in the survey, size of the EMM target population from which the EMM sample was likely to have been drawn, indication of whether a majority subgroup was included in the survey, indication of whether the survey was designed as a general population survey
4. Sampling Method	Sampling strategy used for the survey (closed and open responses), sample design of the survey, sampling frame of the survey, sampling units of the survey
5. Sample Size for	Gross/issued sample for the survey as a whole; net/achieved sample for the survey as

download and reuse on the ETHMIGSURVEYDATA website: http://ethmigsurveydatahub.eu/wp-content/uploads/2020/01/FINAL-ETHMIGSURVEYDATA-Metadata-Template-updated-typos-and-new-vars-Jan2020.xlsx [1 October 2020]

the Overall Survey	a whole; overall response rate of the survey, including the response rate calculation used; issues/challenges with the sample identified by the survey producer; details of the survey weights (if used)
6. Sample Sizes for any Subgroups in which the Survey is Partitioned	For each partitioned subgroup of the survey (at least up to five subgroups), gross/issued sample, net/achieved sample, response rate, calculation used for the response rate, issues/challenges with the sample identified by the survey producer
7. Data Collection Information	Name of the individual or entity that undertook the survey fieldwork; data collection mode used; for surveys that included personal interviews, specifics on who conducted the interviews and indication of whether the interviewer spoke any of the EMM target population languages and if so, which ones; indication of whether the questionnaire was offered in the languages spoken by the EMM target population and if so, which ones; average duration/length of the interviews; number of questions included in the survey
8. Availability to Research Community	Availability of the survey dataset, technical documentation, and questionnaire; data archive/repository IDs and DOIs for the survey dataset, technical documentation, and questionnaire; languages in which the survey dataset, technical documentation, and questionnaire are available; data documentation standards used for the survey's technical documentation; details about any access/use restrictions for the survey's dataset
9. Data Producers, Owners, Distributors and Citations	Names of the survey producer, owner, and distributor; contact details for any queries or requests about the survey; recommended citations for the survey dataset, technical documentation, questionnaire, and other related documents
10. Additional Information	Overall quality rating of the survey (for internal use only), comments about the quality rating of the survey, sources of information used to compile the survey's metadata, any other comments about the survey and the metadata compiled

NOTE 1: This table is an adaptation of the one found in Morales et al. (2020a).
NOTE 2: For more detailed information about each section and the information it captures, please consult pages 8-22 of the guidelines document, FINAL CA16111 ETHMIGSURVEYDATA WG1 WG2 Metadata Template, used for the metadata compilation work: http://ethmigsurveydatahub.eu/wp-content/uploads/2020/01/FINAL-CA16111-ETHMIGSURVEYDATA-WG1-WG2-Metadata-Template.pdf; [1 October 2020]

A subset of the 200 plus members of ETHMIGSURVEYDATA (and participants of FAIRETHMIGQUANT for France) were ultimately tasked with searching for existing surveys in their respective country[10] using detailed

[10] As mentioned in footnote 3, the countries contributing to the EMM Survey Registry are currently those that have formally joined ETHMIGSURVEYDATA (n=35).

inclusion/exclusion criteria[11] developed by the data community. They were also asked to compile metadata for each detected and subsequently included survey by filling out the aforementioned Excel-based template. Given that the metadata schema includes over 200 variables, the ETHMIGSURVEYDATA members and the FAIRETHMIGQUANT participants were provided with a detailed guidelines document[12] describing how to code each variable (including how to handle variables for which no information can be provided[13]). Moreover, a rigorous and multi-stage quality check process[14] was adopted by the Task 9.2 team to ensure metadata for each survey was compiled using the provided instructions. Only after a record describing a survey has successfully completed the full quality check process would it undergo technical adaptations so it could be added to the EMM Survey Registry.

All in all, the entire process that starts with searching for surveys and ends with adding their records to the EMM Survey Registry is time, resource, and labour intensive. While this certainly presents an on-going challenge in terms of quickly offering new records on the EMM Survey Registry, it has ensured that the rich and meaningful metadata that was carefully constructed by the data community is properly being used and applied. Ultimately what this means is that visitors to the EMM Survey Registry are provided with the peace of mind that they can rely on and use the information they are viewing about each survey.

2.2.2 Centering the Front-End on the User

The primary user-facing interface of the EMM Survey Registry is the front-end, i.e. the part of the registry that allows users to explore the compiled survey-level metadata. As the front-end needed to be set up so it would be appealing and relevant to a variety of users, the data community recruited prospective users of the EMM Survey Registry from various sectors, disciplines, etc. to provide detailed input, as well as conduct rigorous testing during each stage of the front-end development process. The data community also used the services of Youngminds[15], an IT company with extensive experience creating user-friendly and user-centric tools for

[11] Surveys were only included if the following conditions were met: 1. it was quantitative and sample-based, 2. it was conducted since January 2000 in one of the 30 plus countries formally participating in ETHMIGSURVEYDATA, 3. it focused on at least one dimension of EMM integration, and 4. it met certain EMM sample size criteria.

[12] The guidelines document was produced with support from the SSHOC project funds allocated to the Task 9.2 team. It is now publicly available for download and reuse on the ETHMIGSURVEYDATA website: http://ethmigsurveydatahub.eu/wp-content/uploads/2020/01/FINAL-CA16111-ETHMIGSURVEYDATA-WG1-WG2-Metadata-Template.pdf; [1 October 2020]

[13] As long as there was enough information to confirm that a survey met the inclusion criteria for the EMM Survey Registry, it was included. Therefore, the EMM Survey Registry includes surveys where their respective records have many variables with missing or no information (for an example, see the record about the survey, Legal and equal on the labour market for Roma communities: https://registry.ethmigsurveydatahub.eu/surveys/2215; [1 October 2020]).

[14] The quality control process is detailed in a guidelines document produced by the Task 9.2 team using funds it had been allocated via the SSHOC project. The guidelines document is used by the Task 9.2 team only to conduct quality checks of the compiled survey-level metadata. The most recent version of the guidelines document (last updated on 31 August 2020) is available for download and reuse on the ETHMIGSURVEYDATA website: https://ethmigsurveydatahub.eu/guidelines-document_-metadata-template-version-control-2/; [1 October 2020]

[15] The official website of Youngminds: http://www.youngminds.ro/; [1 October 2020]

ethnic and migration studies projects, to execute the IT aspects of creating the EMM Survey Registry (which includes the front-end). The resulting front-end (currently available in 'beta' version) is, therefore, a response to a myriad of user-made requests to make the EMM Survey Registry a tool for them.

The current front-end version offers users the following functionalities (see images 1-3 below for visual representations of the functionalities):

- **Keyword search:** Users are able to look for specific surveys using a keyword search based on Boolean logic. Users are also provided with a short tutorial about Boolean logic, so they can adeptly leverage this functionality. The keyword search is positioned towards the top of the EMM Survey Registry home page (above the sorting, advanced filtering, and summary of applied filters).
- **Simple filtering:** Users are able to refine their search results using 15 different filters (i.e. metadata variables from the EMM Survey Registry metadata schema). These filters—which focus on information such as when and where the survey was conducted, who the EMM respondents of the survey were, how the survey was conducted, and how accessible the survey is for reuse—were strategically selected based on input from the user testers. The simple filtering is presented on the left-hand side of the EMM Survey Registry home page.
- **Advanced filtering:** Users, via a pop-up window, are able to refine their search results using 28 different filters (the same 15 from the simple filtering plus 13 new filters). As with the simple filtering, the selection of the filters was based on repeat user input and testing. The main difference between the simple and advanced filtering is that with the latter, users can search for specific topics covered in the survey and more EMM-related classifications. The advanced filtering is currently set up with two access points: one as a link next to the simple filtering and one as a link towards the top of the EMM Survey Registry home page (below the sorting).
- **Summary of applied filters:** After a user applies all of the filters (simple or advanced) they want to use a summary of their selection is displayed above the list of surveys. This functionality, which is presented right above the list of surveys, is particularly useful when users employ numerous filters, as it summarizes how the resulting list of surveys will be refined.
- **Sorting:** Users are able to sort the list of surveys based on country, scope (whether the survey was conducted at the national or subnational/local level), region, start/ends dates, EMM target population, and sample size (achieved). These sorting parameters were selected based on what users expressed, as well as what established data archives and repositories (e.g. GESIS's data catalogue) were using for their respective sorting functionality. This sorting functionality is presented towards the top of the EMM Survey Registry home page (below the keyword search).
- **Abbreviated presentation of a record describing a survey:** When viewing the list of surveys, users are presented with an abbreviated presentation of each record about a survey. This abbreviated presentation is intended to provide users with the highlights or critical information about a survey, so they can quickly decide whether or not it is of interest to them.
- **Full presentation of a record describing a survey:** Each survey that is captured on the EMM Survey Registry is set up with its own page that provides the full record about a survey (i.e. all the information that has been compiled for the survey-in-question using the EMM Survey Registry metadata schema).

As the full record contains an immense amount of information, it is structured so that users can easily navigate through the different sections.

It should be noted that future versions of the front-end will not change drastically from the current 'beta' version; instead, they will likely incorporate small or cosmetic updates that further enhance the user experience. The main rationale for this is that the functionalities that are presently available are those that have been thoroughly tested and vetted and have been deemed sufficient for the different prospective users of the EMM Survey Registry.

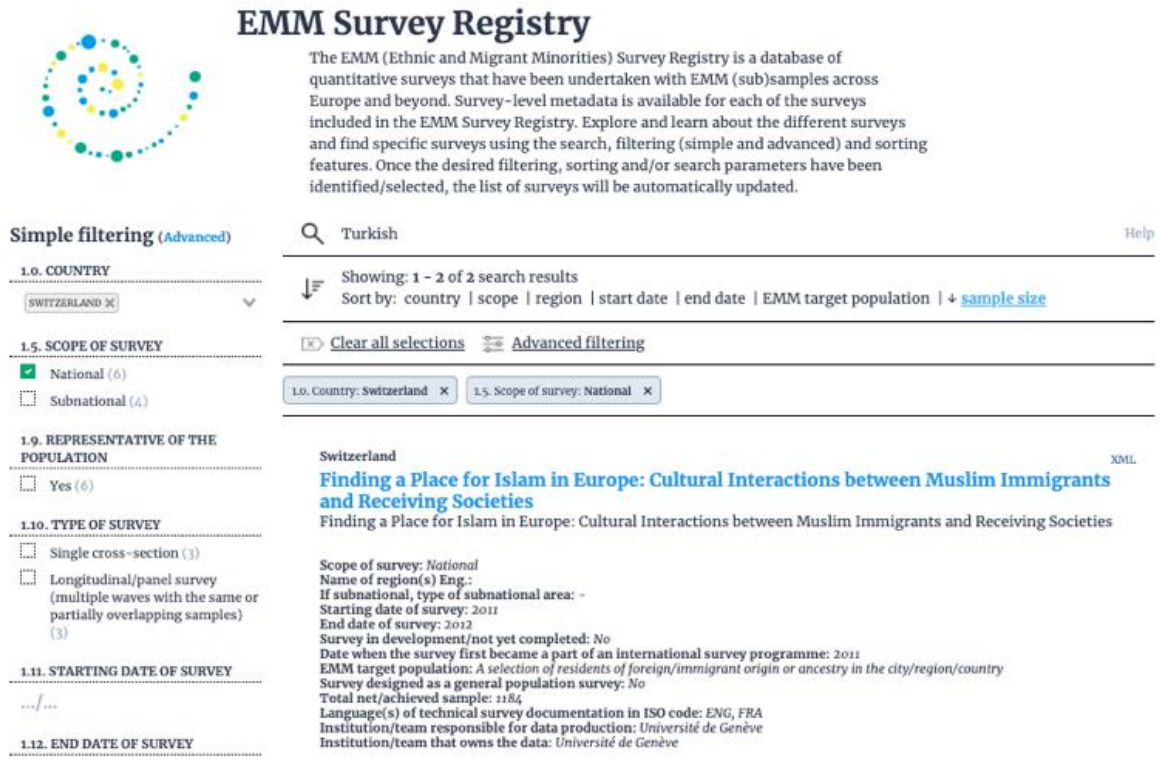


Image 1: *The Front-End Interface: Searching for Swiss National-Level Surveys with Turkish Respondents*

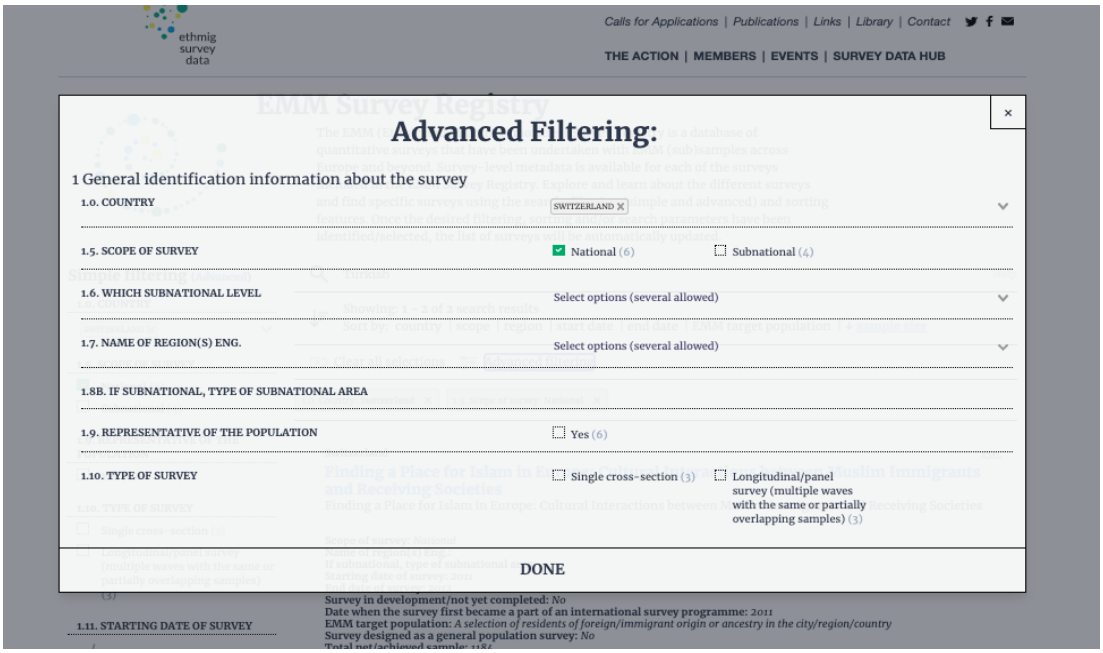


Image 2: Set-Up of the Advanced Filtering



Image 3: Display and Layout of a Full Record of a Survey

2.2.3 Managing the EMM Survey Registry through the Back-End

The back-end of the EMM Survey Registry is where all of its information is stored, organised, and controlled. Similarly to the front-end, the back-end was designed and developed to meet the needs of its targeted-user groups: **a.** the Task 9.2 team who serves as the 'administrators' of the EMM Survey Registry and **b.** data producers who are interested in adding information (using the established EMM Survey Registry metadata schema) about their EMM survey(s).

For group **a**, their identified needs were to be able to edit an existing record of a survey; delete an existing record of a survey[16]; add a new record for a survey that also conforms with the EMM Survey Registry metadata schema; manage back-end users; and manage the variables of the EMM Survey Registry metadata schema, which include adding or removing variables and changing how certain variables are leveraged on the front-end (such as making a variable part of simple or advanced filtering). As such, 'administrator' user permissions were set up for this user group, so they would be able to perform these aforementioned needs[17], particularly to ensure that the EMM Survey Registry is able to adapt to new and/or changing EMM survey research interests (e.g. being able to quickly identify surveys undertaken during the current covid-19 pandemic). Moreover, the interface for the administrative user levels was configured to be user friendly and intuitive, a result of repeated testing conducted by the Task 9.2 team and select members of ETHMIGSURVEYDATA (see images 4-8 for visual representations of the group **a** functionalities).

For group **b**, their only identified needs were to be able to add a new record for a survey and edit a record of a survey that they themselves originally produced (see images 7-8 for visual representations of the group **b** functionalities). Therefore, an 'editor' user permission was established so that they could easily execute these functions, but with careful oversight from group **a**. The oversight from group **a** is such that any new records or any new edits that are made to a record are carefully reviewed, following the same quality check protocols used when metadata was being compiled via the Excel-based template.

It is important to note that at the time this report is being drafted, the back-end has not been made accessible to group **b** users. This was intentionally done so that the Task 9.2 team (with support from select ETHMIGSURVEYDATA members) could sufficiently test that the 'beta' version of the EMM Survey Registry could handle its rapidly expanding 'census' of existing surveys. As the 'beta' version is operating as anticipated, the Task 9.2 team hopes to have the back-end open to group **b** users in fall-winter 2020. Once this is achieved, the Task 9.2 team will explore options to obtain real time feedback, subsequently creating a feedback loop that would inform future developments.

[16] Deleting a record about a survey would occur for one of two reasons: **1.** a record was accidentally created by some user or **2.** a request is made by the data owner, producer, and/or distributor of the survey to remove the record from the EMM Survey Registry.

[17] Two administrative user levels were created: 'super administrator' and 'regular administrator'. They are very similar in what they are permitted to do in the back-end, with the only exceptions being that only 'super administrators' are able to manage back-end users and the EMM Survey Registry's metadata schema.

Finally, users of both groups **a** and **b** will be provided with a unique account that is tied to a valid email address. All users will be able to manage their user information (e.g. changing their email address or password) and will be instructed to not share their account with others for security and safety reasons.

The back-end welcome page for 'administrators', with access points for managing the metadata schema (i.e. Fields), the records for each of the surveys (i.e. Surveys), and back-end users (i.e. Users).

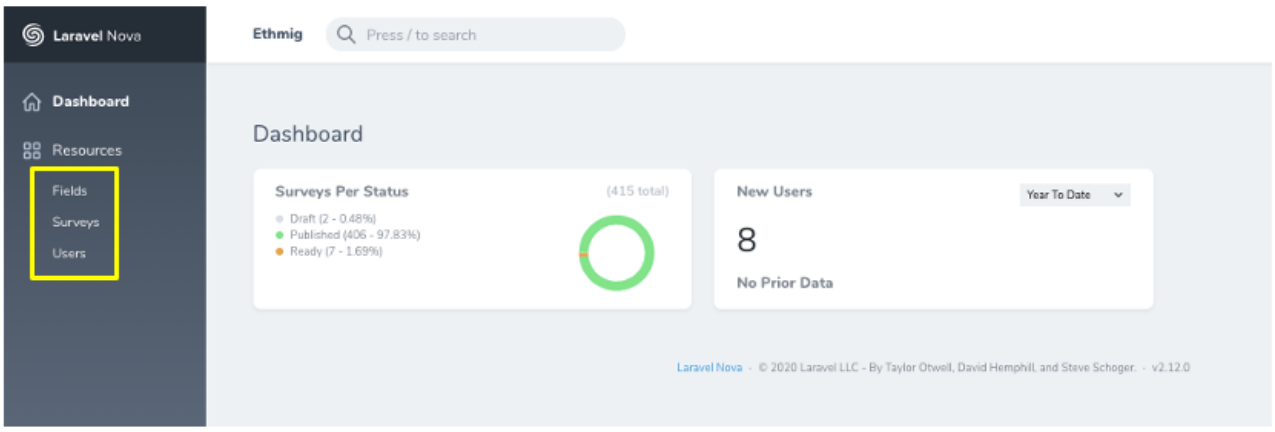


Image 4: *'Administrator' Welcome Page*

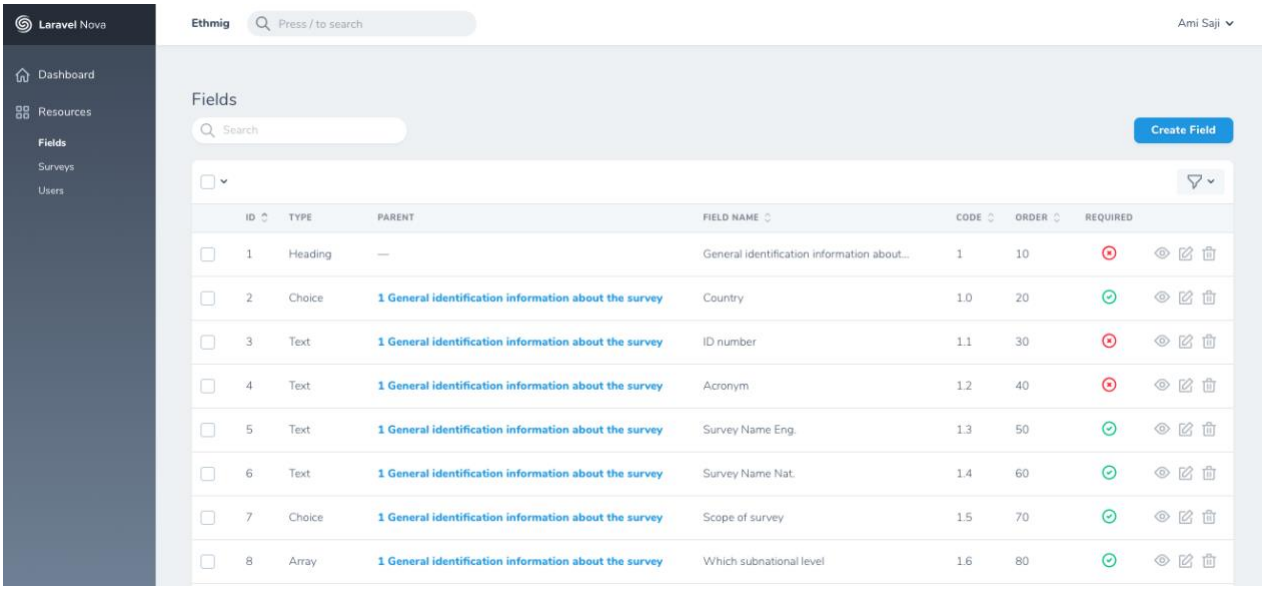


Image 5: *Dedicated Space for the EMM Survey Registry's Metadata Schema*

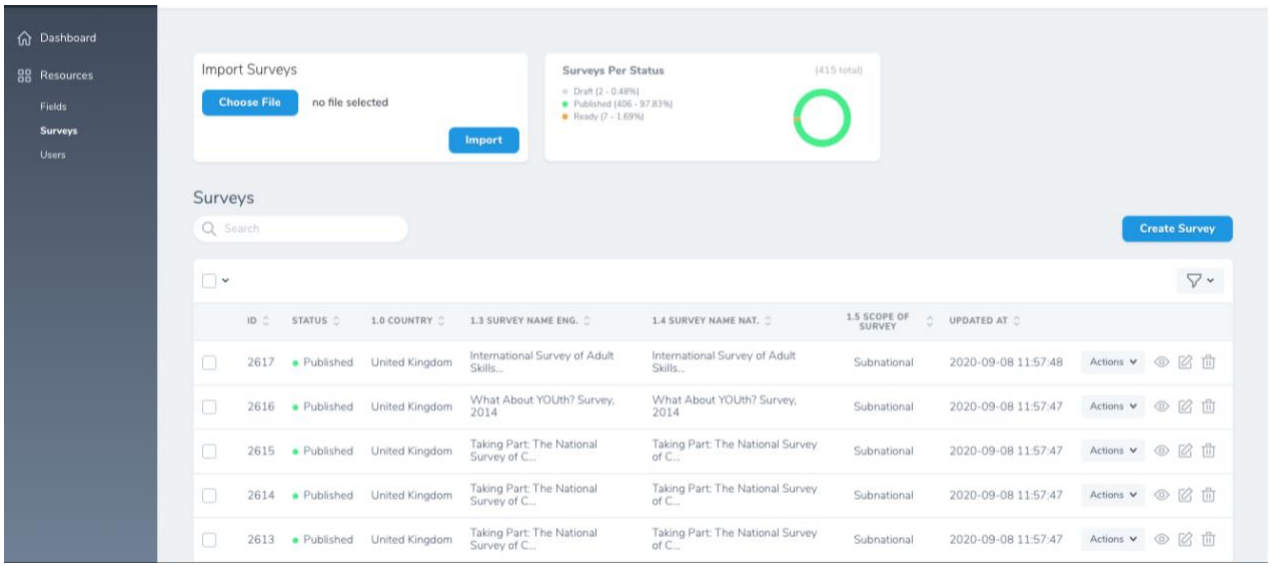


Image 6: *Dedicated Space for the Records of the Surveys*

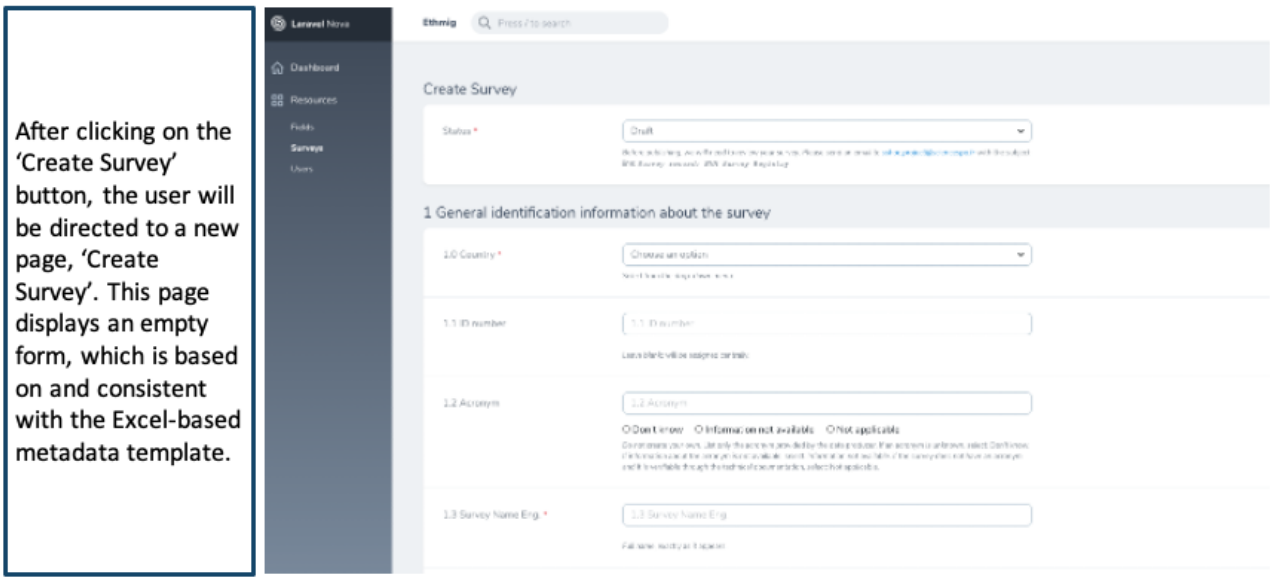


Image 7: *Creating a New Record of a Survey*

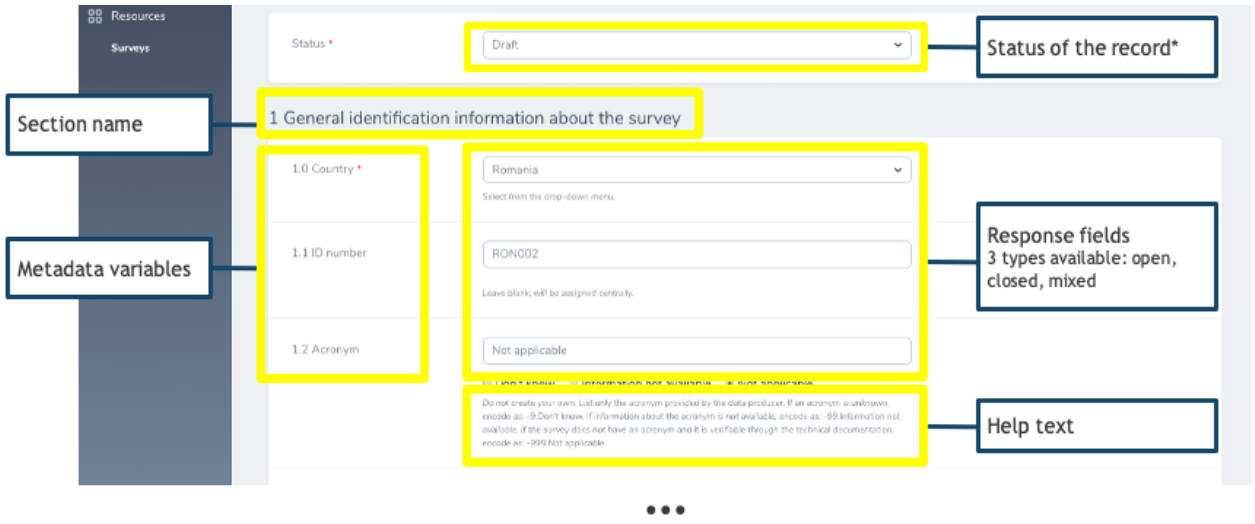


Image 8: *Back-End Form for Creating or Editing a Record of a Survey*

2.3 Adopting DDI Codebook

By offering a single point of entry to access rich and meaningful survey-level metadata for existing quantitative EMM surveys in Europe, the EMM Survey Registry helps to make these surveys easier to find. However, in order to help make the EMM Survey Registry's metadata accessible, reusable, and interoperable, they need to be documented using an established and widely used documentation standard for describing surveys in the social sciences, i.e. the research domain that includes ethnic and migration studies. Through numerous consultations with data storage and management experts (e.g. CESSDA members, SSHOC partners, the CDSP at Sciences Po) and analyses of existing data archives or repositories housing surveys from the social sciences, the data community determined that DDI Codebook would be the most appropriate data documentation standard to use, as it has been designed to allow simple documentation of survey data for the purposes of information exchange or archiving (DDI Alliance, 2020; RDA Metadata Standards Directory, 2020).

To adopt DDI Codebook for the EMM Survey Registry's metadata, the data community sought continued guidance from data storage and management experts. With their support, the data community established the following work plan, which is still in progress:

1. **Design a metadata schema for the EMM Survey Registry that can be mapped onto DDI Codebook (COMPLETED):** Through rigorous and extensive testing and trialling by the data community, a metadata schema for the EMM Survey Registry was developed. Each iteration of the metadata schema was also checked, most notably from ETHMIGSURVEYDATA members affiliated with GESIS and CDSP, to ensure the included or selected metadata variables could be appropriately and easily mapped onto DDI Codebook.

2. **Map the finalised metadata schema for the EMM Survey Registry to specific tags from DDI Codebook (COMPLETED):** Once a final version of the metadata schema was established, the Task 9.2 team collaborated with Alexandre Mairot (formerly with the CDSP, a research centre in Sciences Po participating in WP4) and Alina Danciu (currently employed as a quantitative data specialist with the CDSP) to map its metadata variables to specific tags from DDI Codebook. This mapping exercise was completed by Mairot and Danciu, with regular input and discussions with the Task 9.2 team. At its conclusion, the mapping exercise produced: a prototype[18] of a DDI Codebook-compliant XML file and an Excel-based file[19] that shows the exact pairings made between the metadata variables for the metadata schema and the DDI Codebook tags.
3. **Create a functionality within the EMM Survey Registry that produces an XML file for each 'live' record about a survey (i.e. the actual metadata being displayed on the registry for a single survey) in DDI Codebook (COMPLETED):** The IT company identified to implement the programming of the EMM Survey Registry, Youngminds, was provided with the 2 aforementioned outputs of the mapping exercise undertaken by Mairot and Danciu and offered consultations (on an as needed basis) by Mairot and Danciu. As such, Youngminds was able to successfully set up a functionality within the EMM Survey Registry (see image 9) that allows users to download an XML file version of a 'live' record that is presented in DDI Codebook. Youngminds also provided the source code for this functionality to the data community so that it can be adapted and built on for future versions of the EMM Survey Registry.
4. **Develop an API that conforms with OAI-PMH by leveraging the functionality that produces DDI Codebook-compliant XML files (IN PROGRESS):** For machines to harvest or read the DDI Codebook-compliant XML files generated from the EMM Survey Registry, they need to be made accessible through an API. More specifically, the API needs to be set up to conform with OAI-PMH, which is a standard used by data archives and repositories to promote interoperability. Currently, the Task 9.2 team is liaising with members of SSHOC, CESSDA, and ETHMIGSURVEYDATA to find IT professionals who are familiar with OAI-PMH and can programme the API.

As evidenced above, steps 1-3 of the workflow have successfully been executed and completed. While the current covid-19 situation has presented setbacks in pursuing step 4, the Task 9.2 team, on behalf of the data community, is actively trying to develop the API. Nevertheless, the current 'beta' version of the EMM Survey Registry is promoting accessibility, interoperability, and reusability of the metadata it houses, as the registry

[18] The prototype of a DDI Codebook-compliant XML file was produced by Mairot and Danciu: https://drive.google.com/file/d/1RW8Q51ISSDbWLw0lbDdSW9KwyeoEXdzX/view?usp=sharing; [2 October 2020]. The prototype was provided by Mairot and Danciu to the Task 9.2 team, so that the data community (to which the Task 9.2 team belongs) could use it for the purposes of setting up the EMM Survey Registry. Other uses of this prototype will require permission from Mairot and Danciu.]

[19] The Excel-based file that shows the exact mapping of the EMM Survey Registry metadata variables to DDI Codebook tags was also produced by Mairot and Danciu: https://drive.google.com/file/d/1U95tNksGy6bsfiPl74pbmW1e1fCHs9Us/view?usp=sharing; [2 October 2020]. This file was provided by Mairot and Danciu to the Task 9.2 team, so that the data community could use it for the purposes of setting up the EMM Survey Registry. Other uses of this prototype will require permission from Mairot and Danciu.

allows users to obtain XML files that present the 'live' records (for each survey) in a DDI Codebook-compliant manner.



Image 9: *Generating a DDI Codebook-compliant XML File*

2.4 Setting up the EMM Survey Registry to be a Sustainable Tool

The EMM Survey Registry is intended to serve as a useful and relevant tool for the short and long-term. Therefore, it has been designed and developed to promote sustainability through three distinct pathways: **a.** setting up the metadata schema (both conceptually and on the EMM Survey Registry), so that it can easily be adapted and adjusted to new demands and developments in research and policymaking; **b.** designing the front and back-ends of the EMM Survey Registry, so that they can react to new calls to (further) enhance the user experience; and **c.** producing a mechanism, so that new records can be autonomously (though with oversight) added and contributed to the EMM Survey Registry's 'census' of quantitative EMM surveys.

First, in terms of the metadata schema, its structure is one based on sections. Specifically, each section captures a set of related essential information about a survey, such as general identification details or the sampling method (see table 1 from section 2.2.1 for the full list and description of the sections). This means that the metadata schema can easily and logically incorporate new variables, as it would require just inserting the

variables into the relevant section using the back-end (i.e. the interface used for managing all the (meta)data housed on EMM Survey Registry). For example, if a request was made to create a new variable to capture surveys covering the current covid-19 pandemic, it could be added into section 1, as there is a variable dedicated to denoting the main topics of a survey. Then, for any new survey added to the registry, this new variable would automatically be part of the metadata schema that is used for compiling information about the survey; as for all the surveys that are already documented on the registry, a response to this variable could be coded for each one using the back-end (i.e. the editing feature that allows users to change information for a record about a survey). In a similar vein, variables can also be removed from the metadata schema without negatively impacting how information is being displayed about the survey on the front-end; this is because only the variable in question would be deleted from the metadata schema and all the records created for each survey on the EMM Survey Registry.

Second, the front and back-ends are based on a logic that allows instantaneous adjustments to be made on how users are able to navigate all the surveys documented on the EMM Survey Registry. In other words, the sorting, filtering, searching, and viewing (i.e. the short and long form displays of record describing a survey) functionalities that are accessible and visible on the front-end can be modified to meet user needs via the back-end. For instance, if policymakers were interested in using the EMM Survey Registry to find general population surveys with a substantial subsample of EMM respondents[20], the Yes/No variable capturing this information could be added as a filtering option. This would simply be achieved by going to the variable record on the back-end (which is accessed via the designated space for managing the metadata schema) and then checking the boxes for simple and advanced filtering (see image 10 for a visual representation).

Third, as illustrated in sections 2.1-2.3, the EMM Survey Registry has been created to facilitate buy-in from various stakeholders. Even in its current 'beta' version, it is a tool that is not only well aligned with the FAIR principles but is also set up to allow users to easily learn about existing quantitative surveys on EMMs' integration in Europe via the user-friendly and user-centric front-end. Moreover, by being on track to become a 'one stop shop' for information on quantitative EMM surveys, these same users, particularly those who are data producers and/or owners, could be provided with the validation and motivation needed to organically contribute to the EMM Survey Registry. In other words, the EMM Survey Registry is providing the right incentives to the data producers and/or owners to add in new records about their respective survey(s) through the easy to use and access back-end form.

The data community is now working to make this three-pronged sustainability framework fully operational and functional. This means actively pursuing prospective users to showcase the current and anticipated future versions of the EMM Survey Registry. (see Section 3 below for more detailed information about the concerted effort made by the data community to connect and liaise with as many different stakeholders as possible). This

[20] Generally, the general population surveys included in the EMM Survey Registry include an EMM subsample of 400 or more individuals (for countries with relatively small EMM populations) or 800 or more individuals (for countries with relatively large EMM populations). However, some countries have chosen to include general population surveys regardless of the achieved EMM subsample because such surveys were deemed important research for understanding EMMs' integration in their respective country.

also means ensuring that the EMM Survey Registry will have the appropriate administrative or managerial infrastructure (beyond the lifespan of SSHOC, ETHMIGSURVEYDATA, and FAIRETHMIGQUANT) to manage all the (meta)data it houses in the short and long-term.

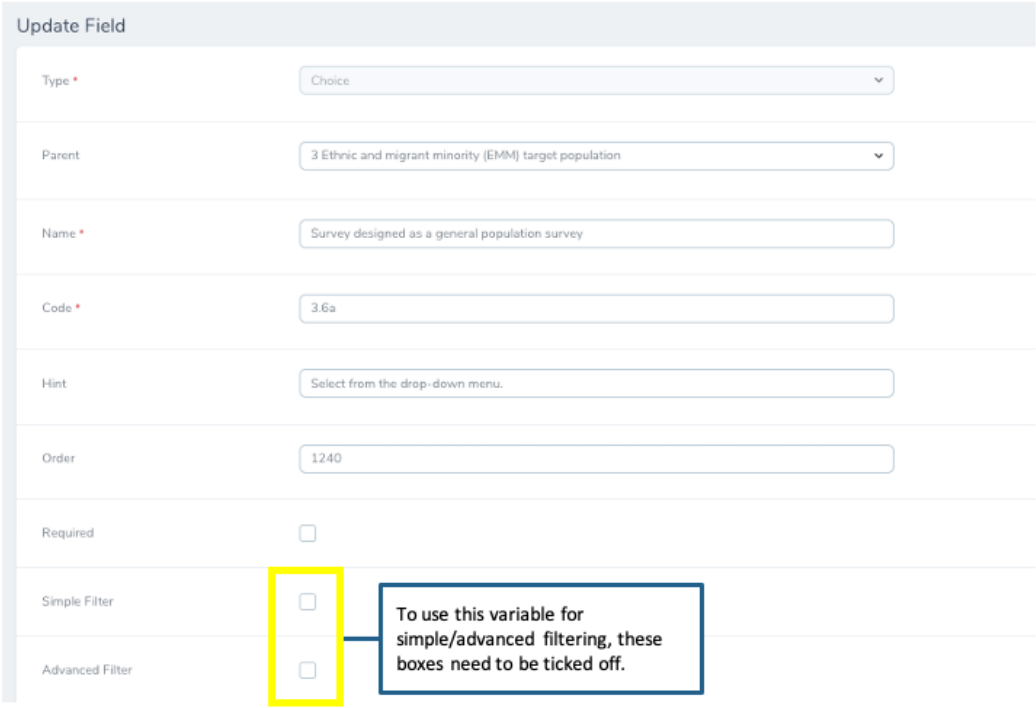


Image 10: *Reconfiguration of a Variable's Display and Use*

3. Connecting and Liaising with Prospective Users of the EMM Survey Registry

The EMM Survey Registry is only valuable and viable if it is leveraged and supported by users of quantitative survey data on EMMs' integration; from the onset of designing and developing the registry, the data community has thus adopted a dissemination and outreach strategy to connect and liaise with a wide-range of prospective users. This strategy is driven by the Task 9.2 team and involves undertaking different types of activities that are tied to different objectives. It also requires that dissemination and outreach be progressively ramped up as the EMM Survey Registry develops and matures.

Table 2 below provides a detailed overview of the key activities already and soon to be delivered by the Task 9.2 team with support from ETHMIGSURVEYDATA and FAIRETHMIGUANT. These activities range from more traditional dissemination and outreach campaigns using social media and other online platforms, to targeted presentations to showcase the EMM Survey Registry as a tool, and to tangible and real world examples of how to use the EMM Survey Registry's metadata (e.g. a report with a preliminary analysis of the metadata, *Surveys to Ethnic and Migrant Minorities across Europe. Identifying Knowledge Strengths and Gaps Using Survey Metadata*[21], which was co-produced by the Task 9.2 team and ETHMIGSURVEYDATA and was targeted towards users of quantitative survey data on EMMs' integration).

Table 2: List of Key Dissemination and Outreach Activities (January 2019 – October 2020)

Activity Type	Activity Description	Date	Location	Activity Owner(s)	Target Audience
Presentation	Provide updates on EMM Survey Registry (alpha version)	July 2019	Rome, IT (COST Action meeting)	T9.2, ETHMIGSURVEY DATA	ETHMIGSURVEYDATA members
Online outreach	Announce launch of alpha version of the EMM Survey Registry	August 2019	SSHOC website	T9.2 (with support from WP2 of SSHOC)	SSHOC project partners; EMM survey data users
Presentation	Inform prospective users about the in-development EMM Survey Registry; promote the importance of adopting the FAIR principles for survey data	February 2020	Bologna, IT (Workshop: International migration data: advances and challenges)	T9.2, ETHMIGSURVEY DATA	EMM survey data users
Presentation	Inform prospective users about the in-development EMM Survey Registry; promote the importance of adopting the FAIR principles for survey data	February 2020	Neuchâtel, CH (NCCR - On the Move Public Lectures Series)	T9.2, ETHMIGSURVEY DATA	EMM survey data users
Social media and online outreach	Announce launch of the alpha (plus) version of the EMM Survey Registry	February 2020	ETHMIGSURVEY DATA Facebook, Twitter, and email lists; SSHOC email list	T9.2 (with support of WP2 of SSHOC), ETHMIGSURVEY DATA	ETHMIGSURVEYDATA members; SSHOC project partners; EMM survey data users
Presentation	Provide updates on EMM Survey Registry (beta version)	March 2020	Brussels, BE (COST Action meeting and Policy Dialogue Conference)	T9.2, ETHMIGSURVEY DATA	ETHMIGSURVEYDATA members; EMM survey data users in Europe
Publication	Share report, *Surveys to Ethnic*	April 2020	Zenodo	T9.2,	ETHMIGSURVEYDATA

[21] This report co-produced by the Task 9.2 team and ETHMIGSURVEYDATA is available on ETHMIGSURVEYDATA's Zenodo page: https://zenodo.org/record/3839677#.Xx7jr5MzZ0s; [2 October 2020]

	and Migrant Minorities across Europe. Identifying Knowledge Strengths and Gaps Using Survey Metadata, produced by T9.2 and ETHMIGSURVEYDATA that analyses metadata currently available on the EMM Survey Registry; provide cogent example of how metadata can be used for research purposes			ETHMIGSURVEY DATA	members; EMM survey data users
Social media and online outreach	Announce launch of the beta version of the EMM Survey Registry	May 2020	ETHMIGSURVEY DATA Facebook, Twitter, and email lists	T9.2, ETHMIGSURVEY DATA	ETHMIGSURVEYDATA members; SSHOC project partners; EMM survey data users
Presentation	Present EMM Survey Registry (beta version) to potential users in France	June 2020	FR (Online meeting for Progedo)	T9.2, ETHMIGSURVEY DATA, FAIRETHMIGQUANT	EMM survey data users in France; data management experts/specialists in France
Online outreach	Inform prospective users in France about the EMM Survey Registry (beta version)	June 2020	Progedo website	T9.2, ETHMIGSURVEY DATA, FAIRETHMIGQUANT	EMM survey data users in France; data management experts/specialists in France
Social media outreach	Announce ETHMIGSURVEYDATA and T9.2's participation in the IMISCOE conference to present a paper based on the survey-level metadata from the EMM Survey Registry and to showcase the EMM Survey Registry itself	July 2020	ETHMIGSURVEY DATA Facebook and Twitter	T9.2, ETHMIGSURVEY DATA	ETHMIGSURVEYDATA members; SSHOC project partners; EMM survey data users
Presentation	Via a panel session, present working paper that analyses survey-level metadata from the EMM Survey Registry	July 2020	2020 IMISCOE Conference - Virtual	T9.2, ETHMIGSURVEY DATA	EMM survey data users; migration researchers
Presentation	Via a roundtable, present EMM Survey Registry (beta version) to potential users	July 2020	2020 IMISCOE Conference - Virtual	T9.2, ETHMIGSURVEY DATA	EMM survey data users; migration researchers
Webinar	Provide a tutorial about the beta version of the EMM Survey Registry and discuss importance of making research data FAIR	October 2020	SSHOC webinar platform	T9.2 (with support from WP2 and WP6 of SSHOC), ETHMIGSURVEY DATA	ETHMIGSURVEYDATA members; SSHOC project partners; EMM survey data users
Presentation	Present working paper that analyses survey-level	October 2020	WAPOR 73rd Annual	T9.2, ETHMIGSURVEY	EMM survey data users; migration researchers;

	metadata from the EMM Survey Registry		Conference and WAPOR LATAM 9th Congress - Virtual	DATA	survey researchers
Presentation	Present EMM Survey Registry (beta version) to potential users in France; present metadata compiled for France	November 2020	Nanterre, FR (Seminar at Paris Nanterre University)	T9.2, ETHMIGSURVEY DATA, FAIRETHMIGQUANT	EMM survey data users in France; data management experts/specialists in France
Presentation	Present how the EMM Survey Registry has adopted DDI Codebook	December 2020	12th Annual European DDI User Conference - Virtual	T9.2, ETHMIGSURVEY DATA	EMM survey data users in France; data management experts/specialists in France

NOTE 1: This table only represents the dissemination and outreach activities undertaken and/or planned at the time that this report was being produced.

4. Conclusion

The Task 9.2 team of the SSHOC project—in close collaboration with ETHMIGSURVEYDATA and FAIRETHMIGQUANT, as part of the wider data community in ethnic and migration studies—has achieved one of its first milestones of setting up its EMM Survey Registry: the online and publicly available database and tool that displays compiled survey-level metadata of existing quantitative surveys undertaken with EMM populations in Europe. This registry, which is an initiative truly driven by a data community from conception to execution, began as an idea generated through ETHMIGSURVEYDATA, an international research network on ethnic and migration studies funded by COST Action 16111. The idea was then successfully transformed into a user-friendly FAIR product through strategic collaborations and detailed planning that involved a network of stakeholders (i.e. data producers, users, distributors, and managers) for the ethnic and migration studies data community (i.e. the Task 9.2 team based at Sciences Po, the ETHMIGSURVEYDATA network, and the FAIRETHMIGQUANT project) throughout the whole development process. Moreover, the EMM Survey Registry is now on track to becoming a sustainable tool, as it is bringing in new users, through deliberate communication and dissemination efforts undertaken by the SSHOC project in coordination with the data community as a whole.

This overall experience of creating the EMM Survey Registry will hopefully serve as a useful and tangible example to the SSHOC project partners, data curation actors, researchers in all sorts of organisations, policymakers, etc., as to how a data community can effectively apply the FAIR principles to their research data and, in turn, produce a novel tool that benefits a wide range of stakeholders. In other words, the EMM Survey Registry development process could be a model that could be upscaled to or even replicated for other data communities that are also multi-disciplinary or in the social sciences. For example, equivalent challenges in finding, accessing and re-using data are confronted by researchers working with health-related surveys, which share many features with those directed at EMM populations: there are many small-scale surveys alongside large national and cross-national surveys that are being produced by researchers from many disciplines (psychologists, biomedical researchers, sociologists, epidemiologists, etc.), and large shares of them are not deposited in data archives or repositories.

Finally, given that the EMM Survey Registry has successfully been set up (albeit in 'beta' version), the Task 9.2 team is confident that by the end of the SSHOC project, the EMM Survey Registry will display information for the 800 plus quantitative EMM surveys detected in the 30 different European countries and will be integrated into the SSHOC Marketplace (and potentially EOSC if selected). Moreover, the Task 9.2 team hopes to capitalise on the achievements made with the EMM Survey Registry and garner more support from a range of stakeholders to ensure the registry's continued progress and success. This means tapping into new resources (particularly funding sources), so that the registry can be adeptly managed and be responsive to new technological, research, and policy demands beyond the lifespan of the SSHOC project, the ETHMIGSURVEYDATA COST Action, and the FAIRETHMIGQUANT project.

5. References

DDI Alliance. (2020). *Getting Started with DDI: Version 2 | Data Documentation Initiative.* DDI Alliance. https://ddialliance.org/resources/getting-started/v2

Huddleston, T., Niessen, J., & Tjaden, J.D. (2013). *Using EU Indicators of Immigrant Integration.* European Commission, Home affairs.

Morales, L., Saji, A., Prandner, D, Bergh, J., Bernat, A., & Méndez Lago, M. (2020a). *Surveys to Ethnic and Migrant Minorities across Europe. Identifying Knowledge Strengths and Gaps Using Survey Metadata. (Version 1).* Zenodo. http://doi.org/10.5281/zenodo.3839677

Morales, L., Saji, A., Prandner, D, Bergh, J., Bernat, A., & Méndez Lago, M. (2020b). *Survey research on the integration of ethnic and migrant minorities (EMMs) across Europe: Identifying research strengths and gaps with the survey metadata of the EMM Survey Registry (First Drat).* 2020 IMISCOE Conference, Virtual.

RDA Metadata Standards Directory. (2020). *DDI (Data Documentation Initiative)*. RDA Metadata Standards Directory GitHub. https://rd-alliance.github.io/metadata-directory/standards/ddi-data-documentation-initiative.html

6. List of Images

Image 1: The Front-End Interface: Searching for Swiss National-Level Surveys with Turkish Respondents

Image 2: Set-Up of the Advanced Filtering

Image 3: Display and Layout of a Full Record of a Survey

Image 4: 'Administrator' Welcome Page

Image 5: Dedicated Space for the EMM Survey Registry's Metadata Schema

Image 6: Dedicated Space for the Records Describing the Surveys

Image 7: Creating a New Record of a Survey

Image 8: Back-End Form for Creating or Editing a Record of a Survey

Image 9: Generating a DDI Codebook-compliant XML File

Image 10: Reconfiguration of a Variable's Display and Use

7. List of Tables

Table 1: Metadata Schema by Section

Table 2: List of Key Dissemination and Outreach Activities